A BRIEF VESTIBULAR DISORIENTATION TEST

Charles S. Harris, Rosalie K. Ambler, and Fred E. Guedry, Jr.

JOINT REPORT

UNITED STATES NAVAL SCHOOL OF AVIATION MEDICINE

AND

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Research Report

A BRIEF VESTIBULAR DISORIENTATION TEST*

Charles S. Harris, Rosalie K. Ambler, and Fred E. Guedry, Jr.

Bureau of Medicine and Surgery
Project MR005.13-6001
Subtask 1 Report No. 82

NASA Order R-47

Approved by

Captain Ashton Graybiel, MC USN
Director of Research

Released by

Captain Clifford P. Phoebus, MC USN
Commanding Officer

1 May 1963

* This research was conducted under the sponsorship of the Office of Life Science Programs, National Aeronautics and Space Administration.

U. S. NAVAL SCHOOL OF AVIATION MEDICINE
U. S. NAVAL AVIATION MEDICAL CENTER
PENSACOLA, FLORIDA

SUMMARY PAGE

THE PROBLEM

The main objective of this study was to determine the reliability and construct validity of a brief vestibular disorientation test (BVD).

FINDINGS

Interrater correlations based on the ratings of 114 naval aviation cadets on the BVD test gave evidence of reliability.

Correlations among average rater judgments, the subjects' self-rate, and the various semantic differential factors provided evidence of construct validity of the BVD test.

ACKNOWLEDGMENTS

The investigators would like to acknowledge with gratitude the services of Mrs. Barbara Martin, Secretary, and Allan Frazier, Hospital Corpsman First Class, USN, who served with two of the authors as raters.

INTRODUCTION

Reactions to unusual maneuvers of the entire body relative to the earth differ markedly from one individual to the next. Some individuals enjoy the experience of tumbling or of free fall, while others are disturbed by identical maneuvers. An illusion of similar maneuvers may be experienced during the Coriolis vestibular reaction which occurs whenever the head is tilted during simple whole-body rotation (1,9). This reaction involves a bizarre sensory influx from the semicircular canals, the otoliths, and proprioceptors which may account for its effectiveness in eliciting nausea or vomiting (4). As in the case of the actual maneuver, some individuals are alarmed, disturbed, and even incapacitated by stimuli which others find amusing and interesting.

Evaluations of individuals exposed to this Coriolis vestibular reaction have been compared with their performance in subsequent flight training in the Netherlands (3,5,6) and Canadian (8) Air Training Commands. Results were encouraging in that the evaluations appeared to predict pilot success and also susceptibility to motion sickness during flight training. Performance in a rotating environment also has been found to be predictive of susceptibility to airsickness and seasickness (2).

The present work represents an assessment of the reliability of this kind of test and introduces a structured rating procedure to permit objective administration by personnel with a modicum of training for the task. In addition, comparisons were made between rater judgments and the subject's evaluation of (a) his own reaction and (b) the meaning of the situation (to the subject) as revealed by the Semantic Differential Technique (SDT) (7).

For a given head movement, intensity of the Coriolis vestibular reaction is directly related to the angular velocity of the rotating device. For a given angular velocity and head movement, the number of such head movements is directly related to probability of producing sickness. Although a severe testing situation may increase ability to discriminate the extremes in the distribution of reactions, it also introduces undesirable aftereffects which, in some cases, curtail the subject's participation in other tests or duties for several hours, a problem encountered by Lansberg (personal communication) when he used an angular velocity of 180 deg/sec (i.e., 30 RPM). To insure that the test and its aftereffects would be brief, the angular velocity used in the present series was 90 deg/sec, and head movements were limited in number.

PROCEDURE

SUBJECTS

One hundred and fourteen naval aviation cadets in their fourteenth week of pre-flight training were used as subjects. They were chosen from nine consecutive classes. About thirteen subjects were randomly selected from each class of approximately thirty-five men. None of the subjects had started training in actual flight, but all had passed a rigorous battery of selection tests.

PROCEDURE ON ROTATING CHAIR

Subjects were instructed in class groups as follows: "We will take you one at a time into another room, where you will make some head movements while you are rotating in a chair. The movements will be made in this manner. (Experimenter demonstrates.) You will experience sensations somewhat similar to those you might experience in an aircraft when diving, climbing, banking, et cetera. Some of you will enjoy the sensations, some of you will be more or less neutral, and some of you will find them unpleasant. We believe a person's performance in this chair may be related to later performance in flight training. The first step, however, is to see if raters can agree on how they rate you on your performance. Therefore, our study is still in the experimental stage, and your performance in the chair will have no bearing on your military career."

Subjects were then taken singly into the experimental room and seated in a rotary chair where practice head movements were made, with the chair stationary, until instructions were clearly understood. The subject was asked to make head movements of 45 degrees in about three seconds with his eyes closed and without mechanical aids. After this, the procedure was as follows: chair accelerated at 15 deg/sec^2 to a constant velocity of 90 deg/sec (15 RPM); after one minute the following positions were assumed by the subject: head right, upright, head left, upright, head right, upright, head left, upright, head forward, upright. Each position was maintained for thirty seconds. Upon completion of this sequence, the chair was stopped by a 15 deg/sec^2 deceleration. The subject was instructed to open his eyes immediately after the sensation of movement stopped.

Four raters, all of whom were inexperienced in this type of task, made independent ratings of each of the 114 subjects. Mimeographed forms were used whereby each rater estimated pallor, sweating, facial expression, speed of recovery, and over-all performance. Included in over-all performance were estimates of the speed and accuracy of head movements, spontaneous comments, intensity of nystagmus observed following deceleration, and behavior upon leaving the chair. Ratings were made for each factor on a ten-point scale, with the lowest point indicating little or no effect and the topmost, strong effect.

Raters were told not to make relative judgments of subjects but to judge each man separately. For example, a rating of 10 on pallor would mean that the subject was extremely pale; a rating of 10 on sweating would mean that the man was sweating profusely. This procedure was adopted to avoid, if possible, the necessity of giving raters a wealth of experience in comparing subjects before they could qualify to administer the test.

QUESTIONNAIRES

The subjects then completed a self-rating scale and the Semantic Differential Scale in a separate room, without knowledge of the raters' judgments. The entire procedure, both the rotation test and filling out the questionnaires, required about ten minutes per subject.

The Self-Rating Scale consisted of six items. Five of the items required the subject to describe by use of a seven-point scale his reactions to the chair in terms of like-dislike, no stomach effects-strong stomach effects, no dizziness-strong dizziness, no sickness feelings-strong sickness feelings, steady on feet-very unsteady on feet. A scale point of one for a given item indicated a favorable or no effect response; a scale point of seven indicated a very unfavorable or strong effect response. The sixth item on this questionnaire was not scaled; it simply asked which head movement bothered the subject the most.

Standard instructions were given for the SDT. Subjects were asked to rate, "My Experience on the Rotating Chair," on sixteen scales of "bipolar" adjectives. Four scales were included to sample the evaluative factor (good-bad, valuable-worthless, fair-unfair, bright-dark), the potency factor (large-small, heavy-light, strong-weak, thick-thin), the activity factor (fast-slow, passive-active, relaxed-tense, cautious-rash), and four scales were chosen to sample what Osgood et al. (7) have said represent "some sort of anxiety factor." These four scales were agitated-calm, cold-hot, awkward-graceful, and constricted-spacious.

SCORING

The raters' judgments on the five factors (pallor, facial expression, sweating, speed of recovery, and over-all performance) were summed for each subject to obtain a score. Since a ten-point scale was used, the range of possible scores was from five to fifty. These scores were used for computing the rater reliability. Means of the four raters' scores were determined for each subject for comparison with the Self-Rating Scale and the Semantic Differential.

The Self-Rating Scale was scored by assigning a numerical value to each of the five scaled items corresponding to the scale position assigned the item by the

subject. These five values were then summed to obtain the score. On this seven-point scale, therefore, the lowest possible score was five and the highest was thirty-five. No evaluation of the response to item number six was attempted for this report.

The Semantic Differential was also scored on a scale from one to seven. Five different scores were obtained for each subject. Mean scores were based on the four scales for each factor [evaluative (E), potency (P), activity (A), and anxiety (X)]. In addition, a score based on the combined mean scores of the evaluative factor and anxiety factor was derived. This E and X score was determined by adding the mean E score to seven * minus the mean X score. For example, if an individual had a mean E score of 4.00 and a mean X score of 3.00, his E and X score would be 4 + (7-3) = 8.00. Thus a high E and X score would be indicative of the subject's finding the test situation enjoyable and not associated with anxiety. Hence it was predicted that a significant negative correlation would be obtained between rater estimates (high score = strong effects) and the E and X score (high score = enjoyable and anxiety-free).

RESULTS

Table I presents the intercorrelations among the four raters in their judgment of the subjects' reactions to head movements in the rotating chair. Also included are intercorrelations between each of these raters and the mean rating based on the four judgments. All of the correlations presented in this table are significant beyond the .001 level of confidence and are of sufficient magnitude to give strong evidence of a reliable measure.

Table II presents further evidence of reliability. Kendall's coefficient of concordance, or $\underline{W}$, was used to determine the extent of agreement of the rater rankings within each of the nine experimental sessions. This statistical technique permits a single index of agreement when several raters are involved (10). There is little doubt from the evidence in Tables I and II that there existed strong agreement among rater observations.

Correlations between rater judgments, self-ratings, and the various factors of the SDT are presented in Table III. Significant correlations were found between all items, but particularly noteworthy were the relatively high correlations between self-rate and raters' estimates and between self-rate and the E and X score.

---

* Maximum score possible was seven.

Table I

The Intercorrelations Obtained from Rater Judgments of the Performance of One Hundred and Fourteen Naval Aviation Cadets on the Rotating Chair

	Rater 1	Rater 2	Rater 3	Rater 4	Average of Four Raters
Rater 1	---	.717	.711	.683	.873
Rater 2		---	.752	.781	.890
Rater 3			---	.684	.920
Rater 4				---	.859
Average of Four Raters					---
$\bar{X}$	13.04	11.68	14.12	12.71	12.92
σ	5.27	3.80	7.92	4.38	4.77

Table II

Measures of Raters' Agreement for Each of the Nine Different Testing Sessions

Group	N	Obtained*	Corrected*
I	13	.67	.92
II	13	.59	.76
III	15	.62	.78
IV	12	.79	.88
V	13	.70	.84
VI	14	.64	.85
VII	11	.61	.69
VIII	13	.79	.98
IX	11	.72	.84

* Coefficient of concordance statistic or Kendall's $\underline{W}$. Correction was required for ties in rankings. All coefficients are significant at better than .001 level.

Table III

Correlations of Average Rater Judgments, Subjects' Self-Rate, and Various Semantic Differential Factors

	Avg. of Raters	Self-Rate	Evaluation (E)	Potency (P)	Activity (A)	Anxiety (X)	E&X
Avg. of Raters	---	.571	-.328	.279	.230	.304	-.362
Self-Rate		---	-.599	.358	.308	.471	-.613
E			---	-.344	-.184	-.439	.859
P				---	.291	.270	-.368
A					---	.375	-.291
X						---	-.814
E&X							---
$\bar{X}$	12.92	11.40	5.11	3.68	3.25	3.35	8.78
σ	4.77	5.50	.98	1.02	.93	.87	1.55

DISCUSSION

The results demonstrated good agreement among four observers in their ratings of individual reactions produced by head movements during whole body rotation at a rate of 90 deg/sec (i.e., 15 RPM). Initial agreement among these raters, who had little instruction and previous experience, was good. Furthermore, as successive groups were evaluated, agreement among raters did not change substantially (implied by Kendall's coefficient of concordance). This suggests that brief training of raters is feasible.

It was further demonstrated that a composite or average of these rater judgments was related to the individual's self-rating and to his perception of the meaning of the situation as indicated by the Semantic Differential Technique. These several interrelations are suggestive of construct validity for this configuration of variables; i.e., two measures presumed to measure the same reaction should be correlated with each other. It is reasonable to expect that the subject's self-rating of his reaction, his assessment of the meaning of the situation, and the rater observations should correlate significantly. Significant correlations were obtained. It was not anticipated that correlations would be high. If in fact the various measures sample partially independent factors, each factor will contribute some unique variance.

Indications of face validity may be adduced from the direction of correlations between the SDT factors and the other measures. Potency of the situation to the individual was positively correlated with rater evaluations and the self-rating. The Evaluative factor, how much the individual liked the situation, was negatively correlated with intensity of reactions indicated by self-rating and rater scores, while Anxiety was positively correlated with these scores. As anticipated, the E and X score was negatively correlated with the same scores.

A further indication of test validity was obtained by testing four subjects prior to an experiment in which they lived in a room which rotated continuously at 10 RPM (Graybiel et al., in preparation) for twelve days. Fortuitously, these four subjects varied markedly in their reaction to this bizarre sensory environment. The order of susceptibility to the ill effects of this bizarre sensory environment was very well predicted by the test.

CONCLUSIONS

1. The Brief Vestibular Disorientation Test is feasible from the point of view of rater reliability.
2. This test is also feasible from the point of view of administration time and subject recovery time (when the 15 RPM rotation rate is used).

3. The intercorrelations among the three measures obtained in the present experiment provide evidence of construct validity for the BVD test.
4. Correlations found between rater judgments, self-ratings, and factors in the SDT may become useful should the test become part of a selection battery.
5. Reliability data confirm the desirability and feasibility of validating test results against the subsequent performance of pilot candidates. A validation study is in progress.

REFERENCES

1. Guedry, F. E., and Montague, E. K., Quantitative evaluation of the vestibular Coriolis reaction. Aerospace Med., 32: 487-500, 1961.

2. Kennedy, R. S., and Graybiel, A., The validity of tests of canal sickness in predicting susceptibility to airsickness and seasickness. Aerospace Med., 33: 935-938, 1962.

3. Lansberg, M. P., Vestibular adroitness test (V.A.T.), Trepaesthetic test and modified Bárány test. Aeromed. Acta, Soesterberg, 3: 247-254, 1954.

4. Lansberg, M. P., On the origin of the unpleasant sensations elicited by head movements during after-sensations. Aeromed. Acta, Soesterberg, 4: 67-72, 1955.

5. Lansberg, M. P., Follow-up report on the vestibular tests in candidate pilots. Aeromed. Acta, Soesterberg, 4: 159-162, 1955.

6. Lansberg, M. P., A Primer of Space Medicine. New York: Elsevier Publishing Co., 1960. Pp. 71-88.

7. Osgood, C. E., Suci, G. J., and Tannenbaum, P. H., The Measurement of Meaning. Urbana, Ill.: University of Illinois Press, 1957. Pp. 1-342.

8. Powell, T. J., Beach, A. M., Smiley, J. R., and Russell, N. C., Successful prediction of air sickness in aircrew trainees. Aerospace Med., 33: 1069-1076, 1962.

9. Schubert, G., Die physioleschen Auswirkingen der Coriolisbeschleungungen bei Flugzeugsteuerung. Z. Hals-Nas.-u. Ohrenheilk, 30: 595-604, 1932.

10. Tate, M. W., and Clelland, R. C., Nonparametric and Shortcut Statistics. Danville, Ill.: Interstate Printers and Publishers, Inc., 1957. Pp. 19-21.